# Onward to Petaflops Computing

David H. Bailey
March 6, 1997

Ref: ACM Communications, vol. 40, no. 6 (June 1997), pg. 90-92

With programs such as the U.S. High Performance Computing and Communications Program (HPCCP), the attention of scientists and engineers worldwide has been focused on the potential of very high performance scientific computing, namely systems that are hundreds or thousands of times more powerful than those typically available in desktop systems at any given point in time. Extending the frontiers of computing in this manner has resulted in remarkable advances, both in computing technology itself and also in the various scientific and engineering disciplines that utilize these systems. On December 16, 1996, a sustained rate of one Tflop/s (also written one teraflops, namely $10^{12}$ floating-point operations per second) was achieved by the "ASCI Red" system, a system employing some 7,000 Intel Pentium Pro processors at Sandia National Laboratory in New Mexico. Now that this long-sought milestone has finally been achieved, one might ask what lies ahead for high-end computing.

Following the custom of marking advances in computing by factors of 1,000, the next major milestone is a sustained rate of one Pflop/s (also written one petaflops, or $10^{15}$ floating-point operations per second). It should be emphasized that we could just as well use the term "peta-ops," since it appears that large scientific systems will be required to perform intensive integer and logical computation in addition to floating-point operations, and completely non-floating-point applications are likely to be important as well. In addition to prodigiously high computational performance, such systems must of necessity feature very large main memories, between ten Tbyte ($10^{13}$ byte) and one Pbyte ($10^{15}$ byte) depending on application, as well as commensurate I/O bandwidth and huge mass storage facilities. The current consensus of scientists who have performed initial studies in this field is that "affordable" petaflops systems may be feasible by the year 2010, assuming that certain key technologies continue to progress at current rates [5].

To get some idea of the scale of these systems, a one Pflop/s computer could dispatch in three seconds a computation that a current desktop workstation would require a full year to perform. One Pbyte of memory could contain the text of approximately one billion books, which is roughly 1,000 times the size of a typical university library. If one were to construct a petaflops system today, even if one were to employ low-cost personal computer components (ignoring for a moment the daunting difficulties of communication and software for such a system), it would cost some 50 billion dollars and would consume some 1,000 megawatts of electric power.

The need for such enormous computing capability is often questioned, but such doubts can be dismissed by a moment's reflection on the history of computing. It is well known that Thomas J. Watson, a founder of IBM, once ventured that there was a worldwide market of only about six computers. Even the legendary Seymour Cray, who recently passed away following a tragic auto accident, designed his Cray-1 system on the premise that there were only about 100 potential customers. In 1980, after the Cray-1 had already achieved significant success, an internal IBM study concluded that there was only a limited market for supercomputers, and as a result IBM delayed its entry into the market.

In stark contrast to these short-sighted projections, some private homes now have more than Watson's predicted six systems. Further, the latest personal computers have more computational power and main memory than the original Cray-1, and enthusiastic users are clamoring for more. When we observe home users becoming annoyed with the "slowness" of their Cray-class computer when running a spreadsheet or graphics application, it is easier to understand how enormous physical simulations could strain the computing power of existing supercomputers.

High-end scientific computers traditionally have been the province of academic and government research laboratories. But in a significant recent development, parallel supercomputers are increasingly being used by professionals in other arenas, including financial analysts in the Wall Street community and marketing analysts in the consumer banking and retailing industry. These developments are certain to increase the demand for future high-end computers.

In short, the demand for state-of-the-art computing power appears insatiable. Thus we may as well start planning now for petaflops systems. Some of the compelling applications anticipated for petaflops computers include the following [6]:

1. Nuclear weapons stewardship.
2. Cryptology and digital signal processing.
3. Satellite data processing.
4. Climate and environmental modeling.
5. 3-D protein molecule reconstructions.
6. Real-time medical imaging.
7. Severe storm forecasting.
8. Design of advanced aircraft.
9. DNA sequence matching.
10. Molecular nanotechnology.
11. Large-scale economic modeling.
12. Intelligent planetary spacecraft.

To elaborate on just a single item, consider 3-D protein molecule reconstructions, also known as the "protein folding problem." In designing a new drug agent, scientists need to examine many protein molecules, each with a specified nucleotide sequence. But at present it is not possible to reliably determine, except by laboratory experiment, the actual 3-D structure of the resulting protein molecule. And without this knowledge, it is not possible to know whether the molecule will have the proper binding sites to be an effective drug agent. Petaflops computers may be powerful enough to do the necessary computations to determine this 3-D structure in a reasonable amount of time. Such a capability could be a powerful new tool for pharmaceutical research.

Some of these anticipated petaflops computer applications will be scaled-up versions of present-day applications, with evolutionary enhancements. Others will consist of integrated simulations of multiple physical effects. Many of these applications will likely employ advanced visualization facilities, such as immersive or remote visualization environments, which are still under development today. But if the history of computing is any guide, a number of exotic new applications will be enabled by petaflops computing technology. These applications may have no clear antecedent in today's scientific computing, and in fact may be only dimly envisioned at the present time.

In spite of such potential, it is not by any means certain that scientific computers produced by private industry will achieve the level of one Pflop/s by 2010. One of the reasons for this conclusion is the recent turmoil in the scientific computing marketplace, which has led computer vendors to cut long-term research in favor of near-term development, and to focus on the more lucrative low- and mid-level systems instead of high-end systems. This phenomenon has been described as the "truncated pyramid" of the current computing marketplace. Computer vendors are not to be faulted for these developments, since they are merely responding to market forces. But it is becoming clear that if the federal government, for instance, wants to have usable petaflops computers for anticipated future missions, then the federal agencies will need to provide a substantial part of the funding for the research and development needed to make them a reality.

There are a number of difficult technical problems that need to be solved in the next few years if we are to achieve the goal of petaflops computers by the year 2010. Indeed, the anticipated difficulties of developing the necessary hardware technology, determining an optimal system architecture, producing reliable system software, devising efficient algorithms, and ultimately programming petaflops systems, present challenges unprecedented in the history of computing [4].

A key issue for these systems is latency management. When citing the breathtaking increases in memory device density during recent years (following Moore's Law), we often fail to note that the access time of these memory devices has not improved very much during this time, nor is there any reason to expect dramatic improvements in the foreseeable future. Thus the gap between processor speed and memory speed is expected to worsen in the future, particularly if exotic technology is used for processors.

Latency can been dealt with by exploiting concurrency, such as in pipelined or multithreaded architectures. The challenge of overcoming latency, coupled with the need to achieve one Pflop/s aggregate sustained performance, will require enormous system concurrency — possibly as many as 1,000,000 processors. Concurrency of this scale is well beyond anything heretofore attempted in high performance computing. Indeed, the coupled challenges of managing latency and extreme concurrency will drive much of the research that needs to be done over the next few years. Some specific research questions that need to be answered include the following:

Hardware

1. Can we produce a usable petaflops system using commercial, off-the-shelf (COTS) hardware components?
2. Is a hybrid hardware technology approach, such as superconducting RSFQ logic [3] with optical interconnect citeJans, superior to a COTS-based design? (RSFQ technology may be able to achieve petaflops power with only 10,000 to 20,000 processors).
3. How can the power consumption for such a system be reduced to acceptable levels?
4. Is a conventional multiple-instruction, multiple-data (MIMD) distributed memory architecture satisfactory, or is some novel architecture superior?
5. What hardware facilities are needed to manage latency and multiple layers of memory hierarchy?
6. What is the best design for mass storage and I/O on such a system?

Software

1. What operating system design can reliably manage tens or hundreds of thousands of processors?
2. Are radically new programming languages needed, or can existing languages be extended?
3. What specific new language constructs will be required?
4. What is the best way to support I/O, debugging and visualization?
5. What software facilities are needed to manage latency and the memory hierarchy?

Algorithms

1. Do there exist latency tolerant variants of known algorithms?
2. How will the operation count, memory requirement, data locality and other characteristics of various algorithms scale on these future systems?
3. Will variations of classical algorithms suffice for key applications, or will completely new algorithms be needed?

Applications

1. Can anticipated petaflops applications be structured to exhibit hundreds of thousands of concurrent threads?
2. What is the best way to implement various applications on proposed system designs?
3. What will be the memory and I/O requirements of future applications?
4. What exotic new applications might be enabled by petaflops systems?

These research questions raise provocative issues about the future of all computing, not just high-end scientific computing. After all, high levels of parallelism are inevitable for all classes of computing, even desktop systems. Thus it is likely that answers to these questions may have impact far beyond the realm of large-scale scientific computing.

There is already a growing research community working on these and related problems of petaflops computing. For example, recently the National Science Foundation awarded a number of research grants to explore system architectures for petaflops computers. These researchers presented their preliminary results at the Frontiers '96 conference [1], which was held in Annapolis, Maryland USA in October 1996. More studies are planned. Onward to petaflops computing!

# References

[1] *Proceedings of Frontiers '96*, IEEE Computer Society, Oct. 1996.

[2] J. Jans, "Planar Packaging of Free-Space Optical Interconnections," *Proceedings of the IEEE*, vol. 82, no. 11 (1994).

[3] S. V. Polonsky, A. F. Kirichenko, V. K. Semenov and K. K. Likharev, "Rapid Single Flux Quantum Random Access Memory," *IEEE Transactions on Applied Superconductivity*, vol. 5, no. 2 (June 1995), pg. 3000-3005.

[4] Thomas Sterling, Paul Messina and Paul Smith, *Enabling Technology for Petaflops Computers*, MIT, Cambridge, 1995.

[5] Thomas Sterling and Ian Foster, "Proceedings of the Petaflops Systems Workshops," Technical Report CACR-133, California Institute of Technology, Oct. 1996.

[6] Rick Stevens, "The 1995 Petaflops Summer Study Workshop," Aug. 1995. This report and some related material is available from http://www.mcs.anl.gov/petaflops.